

DOCUMENT RESUME

ED 480 881

TM 035 266

AUTHOR Linton, Thomas H.; Kester, Donald
TITLE High Stakes Testing in Texas: Does the TAAS Tell the Whole Story?
PUB DATE 2003-04-00
NOTE 10p.; Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, IL, April 21-25, 2003).
PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS
DESCRIPTORS Academic Achievement; *Achievement Gains; *Achievement Tests; *High Stakes Tests; Mathematics Tests; *Norm Referenced Tests; Reading Achievement; Standardized Tests; State Programs; Testing Programs
IDENTIFIERS Texas; *Texas Assessment of Academic Skills

ABSTRACT

This study was conducted to determine whether norm-referenced test data and Texas Assessment of Academic Skills (TAAS) data from 1995 and 1999 showed similar gains in reading and mathematics achievement. TAAS data were collected from all students in Texas who were in the accountability subset of scores used for rating schools, approximately 200,000 students per grade level. Scores from the Metropolitan Achievement Test (MAT-7) were available for approximately 12,500 students for each grade level and each subject area. Effect size comparisons show that TAAS gains were much larger than gains on the MAT-7. The TAAS tests were never designed to measure general academic achievement, but rather whether or not students can perform satisfactorily on a series of objectives measuring basic reading and mathematics skills selected by the Texas Education Agency. Interpretation of TAAS results as a measure of general academic achievement goes well beyond the intended purpose and the psychometric design of the test. (Contains 6 tables, 2 figures, and 11 references.) (SLD)

High Stakes Testing in Texas: Does the TAAS Tell the Whole Story?

PERMISSION TO REPRODUCE AND
DUPLICATE THIS MATERIAL IN ANY FORM,
AND IN ELECTRONIC MEDIA FOR
COLLECTION SUBSCRIPTIONS ONLY,
HAS BEEN GRANTED BY

T. Linton

2A TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☒ This document has been reproduced as
received from the person or organization
originating it.

☐ Minor changes have been made to
improve reproduction quality.

* Points of view or opinions stated in this
document do not necessarily represent
official OERI position or policy.

By
Thomas H. Linton, Ph.D.
Texas A&M University-Corpus Christi

Donald Kester, Ph.D.
Texas A&M University-Corpus Christi

**Presented at the American Educational Research Association annual meeting
Chicago, Illinois**

April 2003

High Stakes Testing in Texas: Does the TAAS Tell the Whole Story?

State-mandated achievement testing, commonly referred to as high stakes testing, has grown significantly in recent years. These tests have become an integral part of standards-based reform and are often used to make educational decisions about tracking, promotion, and graduation of students. The most visible high-stakes testing program is in Texas, where the Texas Assessment of Academic Skills (TAAS) has been administered since 1990.

A number of research studies have reported large, consistent gains on the TAAS and cited these results as evidence that the TAAS has fueled a drive to increase the achievement level of students in Texas (Texas Education Agency, 2000; Hurley, Chamberlain, Slavin, and Madden, 2001; Jerald, 2001).

Few research studies have compared norm-referenced test results with TAAS results. These studies have consistently found either no gain or small achievement gains for both minority and white students, but have been unable to substantiate the large gains reported on the TAAS. Amrein and Berliner, after examining test results for 18 states with high-stakes testing programs, reported that student achievement remained at the same level or went down after the high-stakes testing policies were instituted. Camilli, using a cohort analysis of gains on the National Assessment of Educational Progress (NAEP) math test from grade 4 to grade 8 found that Texas ranked 17th among 35 states. Two studies that directly compared TAAS and NAEP test results (Klein, Hamilton, McCaffrey, & Stecher, 2000; Linton and Kester, 2003), found that TAAS scores increased significantly over a period of four years while NAEP scores did not.

Comparison of TAAS results with norm-referenced test data has been inconsistent, primarily because few districts in Texas use norm-referenced tests as part of their testing program. However, state-wide norm-referenced data are available for 1995, 1996, and 1999 (Texas Education Agency, 2000).

As a part of the state assessment program, the Texas legislature, in 1993 required the Texas Education Agency to obtain national comparative data in the subject areas and grade levels for which the TAAS is administered. The TEA contracted with Harcourt Educational Measurement (HEM) to collect norm-referenced test data in reading and math in grades 3 through 8 and grade 10. The Metropolitan Achievement Test (MAT-7) was administered to a statewide sample of students in 1995, 1996, and 1999. This testing requirement was discontinued after the 1999 test.

Purpose of the Study

The purpose of the study was to determine whether norm-referenced test data and TAAS data from 1995 and 1999 showed similar gains in reading and mathematics achievement.

Research Question

Do students in Texas have similar gains on the TAAS and the MAT-7 in reading and math from 1995 to 1999?

Sample

TAAS data was collected from all students in Texas who were in the accountability subset of scores used for rating schools. The accountability subset of students consisted of all regular education students and nonexempt limited English proficient (LEP) students in grades 3 through 8 and 10. Special education students were not included in the state accountability totals in either 1995 or 1999. The sample of students with TAAS data averaged approximately 200,000 students per grade level.

The sample of students who took the Metropolitan Achievement Test consisted of a stratified cluster sample of all classrooms in Texas in grades 3 through 8 and 10 that were not classified as special education or gifted/talented. Instructions for selecting the sample classrooms were sent to each participating school district by HEM. Each district was responsible for following instructions and identifying classrooms to participate in the study. A sample of approximately 12,500 students was selected for each subject area and grade level for which TAAS data was collected (grades 3-8 and 10).

Research Design

The research design was primarily descriptive, although effect sizes were calculated to determine the magnitude of gains from 1995 to 1999.

Metropolitan Achievement Test (MAT-7) Results

Reading and math scores (in Normal Curve Equivalent scores) for the Metropolitan Achievement Test (MAT-7) for 1995, 1996, and 199 are presented in Tables 1 and 2

Table 1
Metropolitan Achievement Test (MAT-7)
Mean NCE Scores in Reading

Reading	Grade						
	3	4	5	6	7	8	10
Mat-7 1995	48.9	47.2	48.4	47.3	48.4	46.3	43.1
Mat-7 1996	49.9	49.3	50.2	48.2	47.7	47.2	44.5
Mat-7 1999	50.3	52.0	51.2	48.9	49.9	48.4	51.7

Table 2
Metropolitan Achievement Test (MAT-7)
Mean NCE Scores in Math

Math	Grade						
	3	4	5	6	7	8	10
Mat-7 1995	58.0	58.9	58.2	53.3	51.9	52.1	51.3
Mat-7 1996	54.7	56.5	57.1	54.2	53.3	50.1	47.4
Mat-7 1999	55.4	59.4	57.1	54.8	55.1	53.4	54.9

The results indicate that in both reading and math and at every grade level, student mean NCE scores remain relatively constant from 1995 to 1999. The only exception is grade 10 reading, where students exhibited a substantial gain. The test publishers note that comparisons of changes in specific values across years should be made with caution because the national norm sample for the 1995 and 1996 MAT-7 was different than the national norm sample for the 1999 MAT-7. They suggest that longitudinal comparisons should only be used to identify general trends (TEA, 2000).

TAAS Results

The Texas Learning Index (TLI) is a scale score developed specifically for the state of Texas. This scale score is unique to the state of Texas and is not used with any test other than the TAAS. The TLI allows practitioners to compare scores for a grade level/subject from year to year. TLI scores are not comparable between subject areas. Mean reading and math TLI's for 1995 and 1999 TAAS scores are presented in Tables 3 and 4.

Table 3
Texas Assessment of Academic Skills (TAAS)
Mean TLI Scores in Reading

Reading	Grade						
	3	4	5	6	7	8	10
TAAS 1995	77.3	79.5	79.0	79.0	77.9	77.0	77.0
TAAS 1999	83.5	84.8	84.8	84.3	82.0	83.9	84.1

Table 4
Texas Assessment of Academic Skills (TAAS)
Mean TLI Scores in Math

Math	Grade						
	3	4	5	6	7	8	10
TAAS 1995	72.7	73.8	73.8	71.7	70.9	68.8	70.5
TAAS 1999	77.9	80.5	83.0	81.2	80.4	80.0	78.5

Comparison of TAAS reading scores for 1995 and 1999 show a relatively consistent increase for all grade levels tested. Average increase in reading was 5-7 TLI points, with the exception of grade 7. Comparison of TAAS math scores showed a similar pattern of growth, but there was a greater increase. The average increase in math was 7-10 TLI, with the exception of grades 3 and 8.

Effect Size Comparisons

In order to compare gains on the MAT-7 and the TAAS from 1995 to 1999, comparisons were reported in terms of "effect" sizes. The effect size statistic chosen for this study was the *d* statistic. A value of .2, .5, or .8 is generally interpreted as small, medium, or large effect size, respectively (Green and Salkind, 2003). The major advantage of using effect size is that it provides a common metric across tests.

The publisher for the MAT-7 had specifically cautioned that comparing the 1999 MAT-7 results with the 1995 or 1996 MAT-7 results should be used only to identify general trends because the 1999 test had been re-normed. With this caution in mind, the researchers conducted further analyses in order to explore the stability of the MAT-7 effect size. Specifically, additional effect size calculations were done for pair-wise comparisons between 1995 MAT-7 and 1999 MAT-7 mean NCE scores, 1995 MAT-7 and 1996 MAT-7 mean NCE scores and 1996 MAT-7 and 1999 MAT-7 mean NCE scores (Table 5)

Table 5
Effect Size for MAT-7
Reading

MAT-7	Grade						
	3	4	5	6	7	8	10
1995-1999	.066	.023	.138	.076	.071	.100	.408
1996-1999	.019	.128	.047	.033	.104	.057	.342
1995-1996	.047	.100	.085	.043	-.033	.043	.067

Table 6
Effect Size for MAT-7
Math

MAT-7	Grade						
	3	4	5	6	7	8	10
1995-1999	-.123	.024	.052	.071	.152	.062	.166
1996-1999	.033	.138	.000	.028	.085	.157	.342
1995-1996	-.157	-.114	-.052	.043	.066	-.094	-.185

Analysis of the effect size comparisons show a consistent pattern for all pair-wise combinations of test data. Effect size is similar for reading and math at all grade levels with the possible exception of grade 10. With the exception of grade 10, effect sizes are small (d statistic $< .200$). The results indicate a longitudinal trend of small gains or losses on the MAT-7. This is consistent with previous research using NAEP data (Klein, 2000; Linton, 2003). The consistency of the effect sizes found among the three sets of MAT-7 data in the present study suggest that comparison of the effect size for the MAT-7 and the TAAS is appropriate. Comparison of effect size for MAT-7 reading/math and TAAS reading/math is presented in Figures 1 and 2.

Figure 1
MAT and TAAS Reading Tests
Comparison of Effect Size

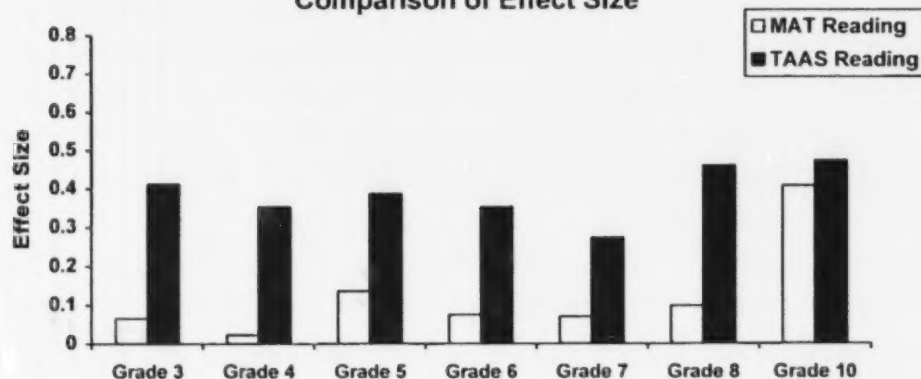


Figure 1 shows that students at every grade level had higher MAT-7 scores in 1999 than in 1995, but with the exception of grade 10, there was only a small gain over the four year period. In contrast, TAAS scores showed a moderate gain over the same four years.

Figure 2
MAT and TAAS Math Tests
Comparison of Effect Size

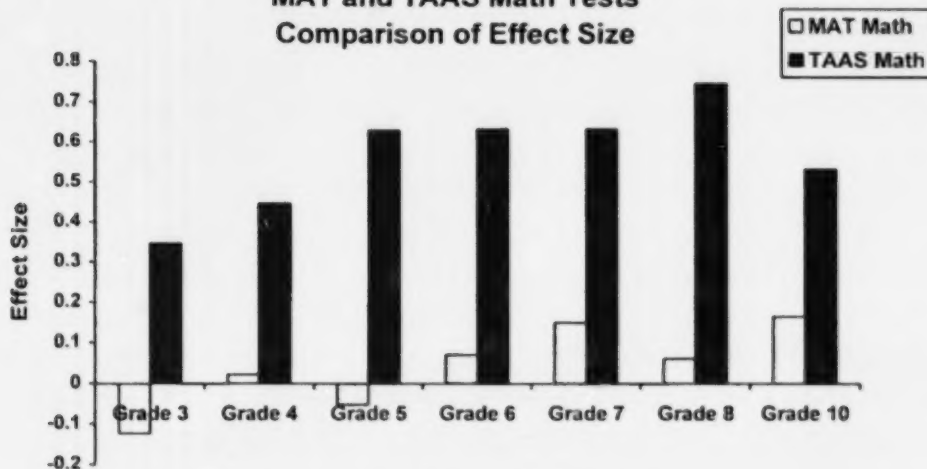


Figure 2 shows that MAT-7 scores increased slightly in grades 4, 6, 7, 8, and 10, but decreased in grades 3 and 5. The gains and losses were very small and could be chance fluctuations. Again, TAAS scores show moderate to large gains from 1995 to 1999. Effect size comparisons show that TAAS gains were much larger than MAT-7 gains.

Discussion

The large gains reported here and in other studies (Hurley, 2001; Jerald, 2000; Klein, et al., 2000; Linton and Kester, 2002) on the TAAS tests, which are minimum skills tests tailored specifically to the Texas reading and mathematics curricula, have not been substantiated by other test results. As previously noted, Klein, et al (2000) in a study of fourth and eighth grade reading and mathematics reported large gains on the TAAS but little or no gain on the National Assessment of Educational Progress (NAEP) over a 4 year period. Linton and Kester (2003), in a study of eighth grade mathematics found large TAAS gains and small NAEP gains from 1996 to 2000.

The results of this study are also consistent with the general pattern of norm-referenced test scores reported by Amrein and Berliner (2002). After examining SAT, ACT, and NAEP test results for 18 states with high-stakes testing programs, they reported that student achievement remained at the same level or went down after the high-stakes testing policies were instituted. In other words, the large gains reported on state-mandated high-stakes tests did not result in gains in achievement as measured by norm-referenced tests of general academic achievement.

Indeed, the TAAS tests were never designed to measure general academic achievement. The TAAS tests are designed to measure whether or not students can perform

satisfactorily on a series of objectives measuring basic reading and mathematics skills selected by the Texas Education Agency. Therefore, the interpretation that large gains on the TAAS represent large achievement gains in a subject area is inappropriate and misleading. In our opinion, interpretation of TAAS results as a measure of general academic achievement goes well beyond the intended purpose and the psychometric design of the test.

The results of this study would seem to have implications beyond the State of Texas. Data from high stakes tests are being used more and more frequently by national, state, and local decision-makers to develop educational policy, formulate curricula, and fund educational initiatives that have far reaching consequences for the nations schools. These decision-makers must be made aware of the psychometric limitations of the tests and avoid over interpretation of the test data. It is vital that decision-makers develop state and national policies based on an accurate picture of what the tests are telling them about the educational achievement of their students.

References

- Amrein, A. L. and Berliner, D. C. (2002). High-stakes testing, uncertainty, and student learning. *Education Policy Analysis Archives*, 10(18). from <http://epaa.asu.edu/epaa/v10n18/>.
- Camilli, G. (2000). Texas gains on NAEP: points of light? *Education Policy Analysis Archives*, 8(42). from <http://epaa.asu.edu/epaa/v8n42/>.
- Green, S. & Salkind, N. (2003). *Using SPSS for windows: analyzing and understanding data* (3rd ed.). Upper Saddle River, NJ: Prentice-Hall.
- Hurley, E., Chamberlain, A., Slavin, R., & Madden, N. (2001). Effects of success for all on TAAS reading scores: a Texas statewide evaluation. *Phi Delta Kappa*, 82(10), 750-756.
- Jerald, C. (2001). Real results, remaining challenges: the story of Texas education reform. *The Education Trust*, (commissioned by The Business Roundtable) [online] Available: (<http://www.brtable.org/document.efm/532>).
- Klein, S., Hamilton, L., McCaffrey, B., & Stecher, B. (2000). *What do test scores in Texas tell us?* *Education Policy Analysis Archives*, 8(49). from <http://epaa.asu.edu/epaa/v10n18/>.
- Linton, T., & Kester, D. (2002). Exploring the achievement gap between white and minority students in Texas: A comparison of the 1996 and 2000 NAEP and TAAS

eighth grade mathematics test results. *Educational Policy Analysis Archives* 11(10). From: <http://epaa.asu.edu/epaa/v8n42/>.

Linton, T., & Kester, D. (2002). Investigating the Achievement Gap Between White and Minority Students in Texas: Should We Be Concerned about a Ceiling Effect on the TAAS? Paper presented at the annual meeting of the American Educational Research Association, New Orleans, La.

Texas Education Agency (2000, April 20). State's exit-level TAAS performance sets another record: 10th graders hit 90 percent mark on two sets of tests. News Release. Austin, TX

Texas Education Agency, Office of Policy Planning and Research. (2000). *The 2000 accountability rating system for Texas public schools and school districts*. Austin, TX.

Texas Education Agency, Student Assessment Division (2000). *1999 Texas Comparative Data Study*. Austin, TX. From <http://www.tea.state.tx.us/student.assessment/resources/studies/ncds>.

Dr. Tom Linton is an associate professor in the Educational Leadership Doctoral Program at Texas A&M University-Corpus Christi. His experience includes 25 years in the Corpus Christi ISD Office of Research Testing, and Evaluation and with various other school districts as an outside consultant and as a project evaluator. He has also served on several advisory committees for the Student Assessment Division of the Texas Education Agency. His e-mail address is: tlinton@falcon.tamucc.edu.

Dr. Don Kester is an associate professor in the Educational Leadership Doctoral Program at Texas A&M University-Corpus Christi. His experience includes 27 years as an administrative consultant in the areas of educational program management, evaluation, research, and student assessment with the Los Angeles County Office of Education. His e-mail address is: don.kester@mail.tamucc.edu.



U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)



REPRODUCTION RELEASE

(Specific Document)

TM035266

I. DOCUMENT IDENTIFICATION:

Title: *HIGH STAKES TESTING IN TEXAS: DOES THE TAAS TELL THE WHOLE STORY?*

Author(s): *THOMAS H. LINTON, DONALD KESTER*

Corporate Source:

Publication Date:

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

Level 1



Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) and paper copy.

The sample sticker shown below will be affixed to all Level 2A documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2A

Level 2A



Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only

The sample sticker shown below will be affixed to all Level 2B documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2B

Level 2B



Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign
here, →
please

Signature: *Thomas H. Linton*
Organization/Address: *TEXAS A+M CORPUS CHRISTI
6300 OCEAN DRIVE, EDC#232
CORPUS CHRISTI, TX 78413*

Printed Name/Position/Title: *THOMAS H. LINTON, ASSOC. PROF*
Telephone: *(361) 825-3364* FAX: *(361) 825-3377*
E-mail Address: *tlinton@falcon.tamucc.edu* Date: *9/5/03*

(Over)

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:

Address:

Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:

Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse: **University of Maryland**
ERIC Clearinghouse on Assessment and Evaluation
1129 Shriver Lab, Bldg 075
College Park, MD 20742
Attn: Acquisitions

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

University of Maryland
ERIC Clearinghouse on Assessment and Evaluation
1129 Shriver Lab, Bldg 075
College Park, MD 20742
Attn: Acquisitions